PartA

November 29, 2024

```
[7]: import csv
     import re
     import pandas as pd
     df = pd.read_csv('nba.csv')
     df = df.fillna('Unknown')
     print(df)
     df.to_csv('nba_modified.csv', index=False)
     df1 = pd.read csv('nba1.csv')
     df1 = df1.fillna('Unknown')
     print(df)
     df1.to_csv('nba1_modified.csv', index=False)
     def validate_data(row):
         column_rules = {
             "Unnamed": lambda x: re.match(r'^-?\d+\$', x) or x == "Unknown",
             "height": lambda x: re.match(r'^-?\d+\$', x) or x == "Unknown",
             "weight": lambda x: re.match(r'^-?\d+\$', x) or x == "Unknown",
             "born": lambda x: re.match(r'^-?\d+\$', x) or x == "Unknown",
             "Player": lambda x: x.isalpha(),
             "collage": lambda x: x.isalpha(),
             "birth_city": lambda x: x.isalpha(),
             "birth_state": lambda x: x.isalpha(),
         }
         for column, rule in column rules.items():
             if column in row:
                 value = row[column].strip()
                 if rule(value) is False:
                     return False
         return True
     with open('Players(processed).csv', 'r') as file:
         reader = csv.DictReader(file)
         is_valid = True
         for i, row in enumerate(reader):
             if i <= 10000:
                 continue
             if not validate_data(row):
                 is valid = False
                 print(f"Invalid data found in row {i}: {row}")
```

if is_valid: print("All data is valid.") G \ Rk Player Age Team Pos 0 1.0 Joel Embiid 28.0 PHI C 66.0 Luka Dončić PG 66.0 1 2.0 23.0 DAL 2 3.0 Damian Lillard 32.0 POR PG 58.0 3 4.0 Shai Gilgeous-Alexander 24.0 OKC PG 68.0 4 63.0 5.0 Giannis Antetokounmpo 28.0 MIL PF . . ••• 536.0 29.0 PF 2.0 675 Deonte Burton SAC 676 537.0 Michael Foster Jr. 20.0 PHI PF 1.0 Frank Jackson 24.0 UTA PG677 538.0 1.0 678 539.0 Alondes Williams 23.0 BRK SG 1.0 679 Unknown League Average Unknown Unknown Unknown Unknown GS MP FGFGA DRB TRB AST 0 66.0 34.6 11.0 20.1 8.4 10.2 4.2 1 66.0 36.2 10.9 22.0 7.8 8.6 8.0 2 58.0 36.3 9.6 20.7 7.3 4.0 4.8 3 20.3 68.0 35.5 10.4 4.0 4.8 5.5 9.6 4 63.0 32.1 11.2 20.3 11.8 5.7 675 0.0 3.0 0.0 1.0 0.0 0.0 0.0 676 0.0 1.0 0.0 0.0 0.0 0.0 0.0 3.0 677 0.0 5.0 0.0 1.0 2.0 1.0 678 0.0 5.0 0.0 0.0 1.0 1.0 0.0 679 Unknown Unknown Unknown Unknown Unknown Unknown Unknown ••• STL BLK TOV PF PTS Awards 0 1.7 MVP-1DPOY-9CPOY-5ASNBA1 1.0 3.4 3.1 33.1 1 32.4 1.4 0.5 3.6 2.5 MVP-8CPOY-8ASNBA1 2 0.9 0.3 1.9 3.3 32.2 CPOY-10ASNBA3 3 1.6 1.0 2.8 2.8 31.4 MVP-5CPOY-7ASNBA1 4 0.8 0.8 3.9 3.1 31.1 MVP-3DPOY-6ASNBA1 675 0.0 0.0 0.0 0.0 0.0 Unknown 676 0.0 0.0 0.0 0.0 0.0 Unknown Unknown 677 0.0 0.0 0.0 0.0 0.0 678 0.0 0.0 2.0 1.0 0.0 Unknown 679 Unknown Unknown Unknown Unknown Unknown Unknown Player-additional 0 embiijo01 doncilu01 1 2 lillada01

break

3

gilgesh01

4	а	antetgi01											
 675	L.												
675 676	burtode02 fostemi02												
677		acksfr01											
678	willial06												
679	•	-9999											
[680	rows x 3	32 columns]										
	Rk			Player		Age	Team		Pos		G	\	
0	1.0	Joel Embiid			2	8.0	PHI	C 66.0			.0		
1	2.0		Dončić	23.0 DAI			PG 66.0						
2	3.0	Damian Lillard				32.0 POR							
3	4.0	Shai Gilgeous-Alexander			2	24.0 OKC			PG 68.0				
4	5.0	Gianni	s Antetok	ounmpo	2	8.0	MIL		PF	63	.0		
 675	 E26 0		Doonto	 Dant on			 GAC	•••		2	. ^		
675 676	536.0	Deonte Burton Michael Foster Jr.				29.0 SAC 20.0 PHI			PF PF		.0		
	537.0 538.0								PG		.0		
677 678	539.0	Frank Jackson Alondes Williams				24.0 UTA 23.0 BRK			SG		.0		
679	Unknown	л	League A		Unkn		Unknown	Unkn		Unkno			
013	OHAHOWH		League A	verage	OHKH	.Owii	IIWOIIIIO	Olikli	OWII	Olikilo	WII		
	GS	MP	FG	FGA			DRB	TRB		AST	\		
0	66.0	34.6	11.0	20.1	. 		8.4	10.2		4.2			
1	66.0	36.2	10.9	22.0			7.8	8.6		8.0			
2	58.0	36.3	9.6	20.7			4.0	4.8		7.3			
3	68.0	35.5	10.4	20.3	·		4.0	4.8		5.5			
4	63.0	32.1	11.2	20.3	·		9.6	11.8		5.7			
	•••	•••		·	•••								
675	0.0	3.0	0.0	1.0			0.0	0.0		0.0			
676	0.0	1.0	0.0	0.0			0.0	0.0		0.0			
677	0.0	5.0	0.0	3.0			1.0	2.0		1.0			
678	0.0	5.0	0.0	0.0			1.0	1.0		0.0			
679	Unknown	Unknown	Unknown	Unknown	l	Unkı	nown Uni	known	Unkr	nown			
	STL	BLK	TOV	PF	,	PTS	3			Awar	ds	\	
0	1.0	1.7	3.4	3.1		33.	1 MVP-1	DPOY-9	CPOY-	-5ASNB	8A1		
1	1.4	0.5	3.6	2.5)	32.4	4	MVP-8	CPOY-	-8ASNB	8A1		
2	0.9	0.3	3.3	1.9)	32.2	2	C	POY-1	LOASNB	SA3		
3	1.6	1.0	2.8	2.8	3	31.4	4	MVP-5CPOY-7ASNBA1					
4	0.8	0.8	3.9	3.1		31.	1	MVP-3DPOY-6ASNBA1					
	•••	•••											
675	0.0	0.0	0.0	0.0)	0.0	0			Unkno	wn		
676	0.0	0.0	0.0	0.0)	0.0	0			Unkno	wn		
677	0.0	0.0	0.0	0.0)	0.0	0			Unkno	wn		
678	0.0	0.0	2.0	1.0		0.0				Unkno			
679	Unknown	Unknown	Unknown	Unknown	. Un	knowi	n			Unkno	wn		

```
Player-additional
            embiijo01
0
1
            doncilu01
2
            lillada01
            gilgesh01
3
            antetgi01
4
. .
            burtode02
675
            fostemi02
676
677
            jacksfr01
678
            willial06
679
                -9999
```

[680 rows x 32 columns] All data is valid.

[]: